EDA AND BUSINESS INSIGHTS

FROM E-COMMERCE DATASET

PRESENTED BY: ISHWARYA C
ISHWARYAC2415@GMAIL.COM

EDA

I performed the following five EDA tasks to understand the dataset:

- **Merged Datasets:** Combined three datasets into a final dataset with 1,000 records and 13 columns, ensuring no missing values.
- **Data Cleaning:** Removed one duplicate row (price) and verified the dataset for consistency.
- **Feature Engineering:** Added a MonthsSinceSignup column to quantify customer lifecycle, enabling insights on retention and segmentation for targeted marketing.
- **Column Removal:** Removed TransactionID, CustomerID, and ProductID to eliminate noise and focus on meaningful features.
- **One-Hot Encoding:** Applied to Region and Category columns to prepare categorical data for machine learning models.
- **Feature Scaling:** Min-Max scaling was applied to numerical features to normalize the data, ensuring all features are within the range [0, 1] for improved model performance, particularly when using distance-based algorithms like clustering.

BUSINESS INSIGHTS

1.SALES PERFORMANCE BY REGION:

- SOUTH AMERICA HAS THE HIGHEST SALES, INDICATING STRONG DEMAND. THE COMPANY SHOULD INVEST MORE IN SOUTH AMERICA TO CAPITALIZE ON THIS GROWTH.

- EUROPE COMES SECOND, SHOWING POTENTIAL FOR IMPROVEMENT. THE COMPANY COULD BOOST SALES IN EUROPE WITH TARGETED STRATEGIES.

- ASIA AND NORTH AMERICA ARE PERFORMING LOWER, SUGGESTING THE COMPANY COULD EXPLORE NEW OPPORTUNITIES OR MARKETING STRATEGIES IN THESE REGIONS.


Sales Performance by Region
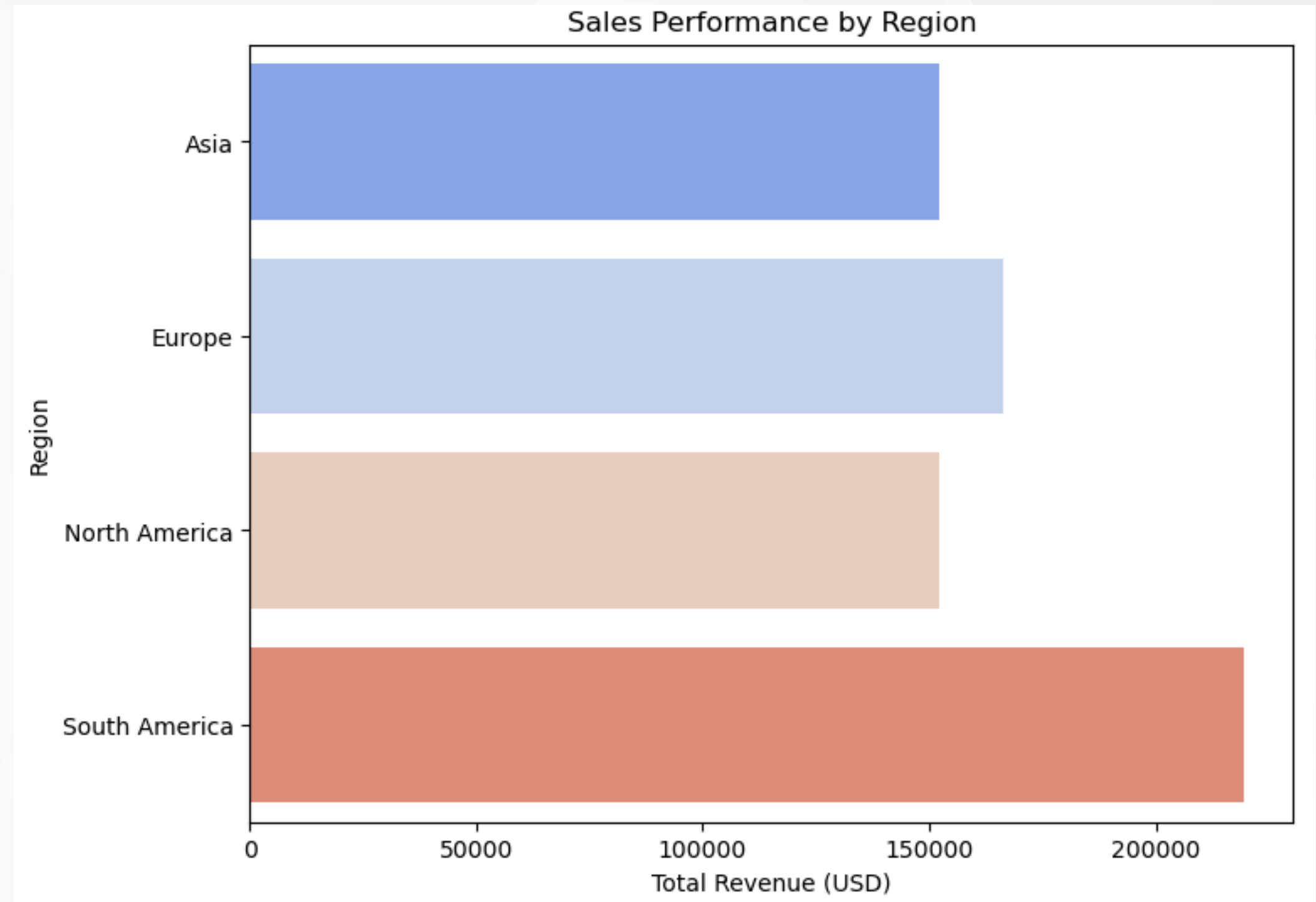
2.CALCULATE QUANTITY SOLD FOR EACH PRODUCT

- TOP 5 PRODUCTS: THESE PRODUCTS ARE THE MOST POPULAR AND IN HIGH DEMAND. FOCUS ON INCREASING STOCK AND PROMOTIONS FOR THESE ITEMS TO MEET CUSTOMER PREFERENCES.

- LEAST 5 PRODUCTS: THESE PRODUCTS HAVE LOW SALES AND MAY BE UNDERPERFORMING. INVESTIGATE REASONS FOR LOW DEMAND, SUCH AS PRICING, AVAILABILITY, OR RELEVANCE, AND CONSIDER PROMOTIONAL STRATEGIES OR PRODUCT IMPROVEMENTS.

- STRATEGY: INCREASE VISIBILITY AND MARKETING FOR TOP PRODUCTS WHILE REEVALUATING OR DISCONTINUING UNDERPERFORMING ITEMS. OPTIMIZE INVENTORY MANAGEMENT BASED ON SALES TRENDS.

5 Most Frequently Purchased Products

5 Underperforming Products
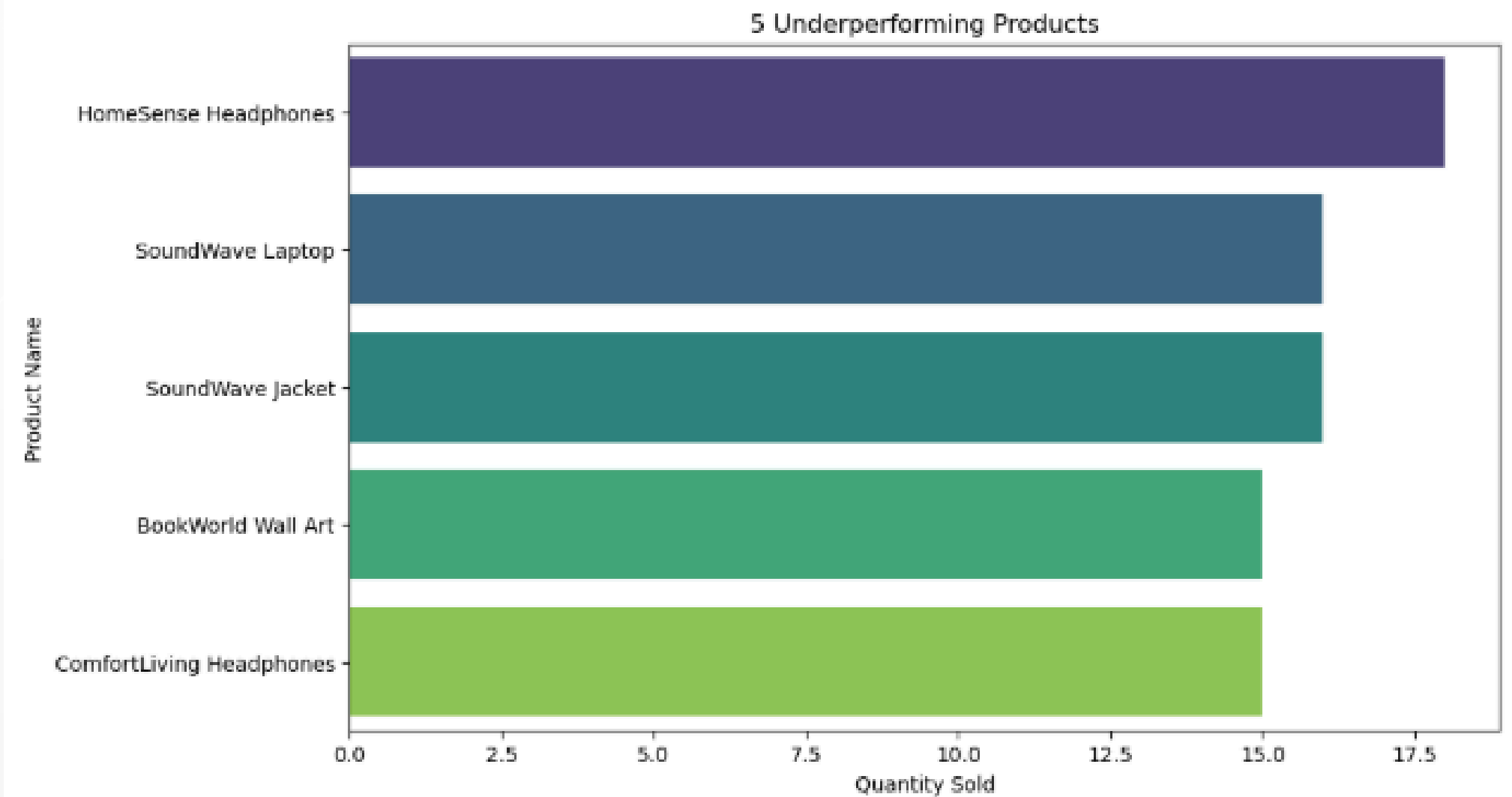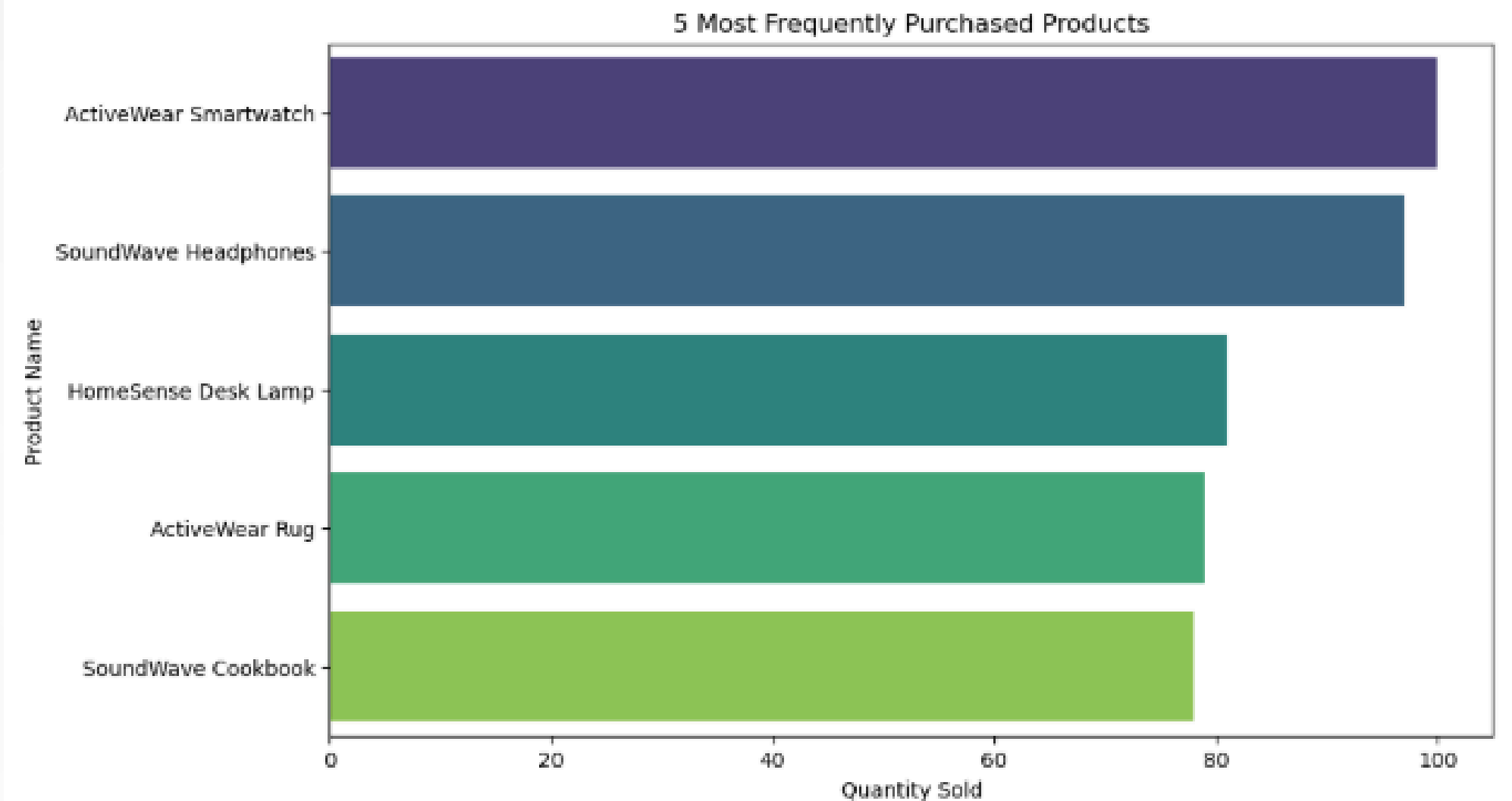
3.TOP 5 HIGH-VALUE CUSTOMERS

- **THE TOP 5 CUSTOMERS, LED BY PAUL PARSONS, ARE THE BIGGEST CONTRIBUTORS TO OVERALL SPENDING.**

- **IDENTIFYING THE REGIONS AND PREFERRED PRODUCTS OF THESE CUSTOMERS WILL HELP TAILOR OFFERINGS.**

- **PRIORITIZING THESE CUSTOMERS FOR LOYALTY PROGRAMS AND PERSONALIZED EXPERIENCES WILL ENHANCE RETENTION AND INCREASE OVERALL SALES.**


Top 5 High-Value Customers
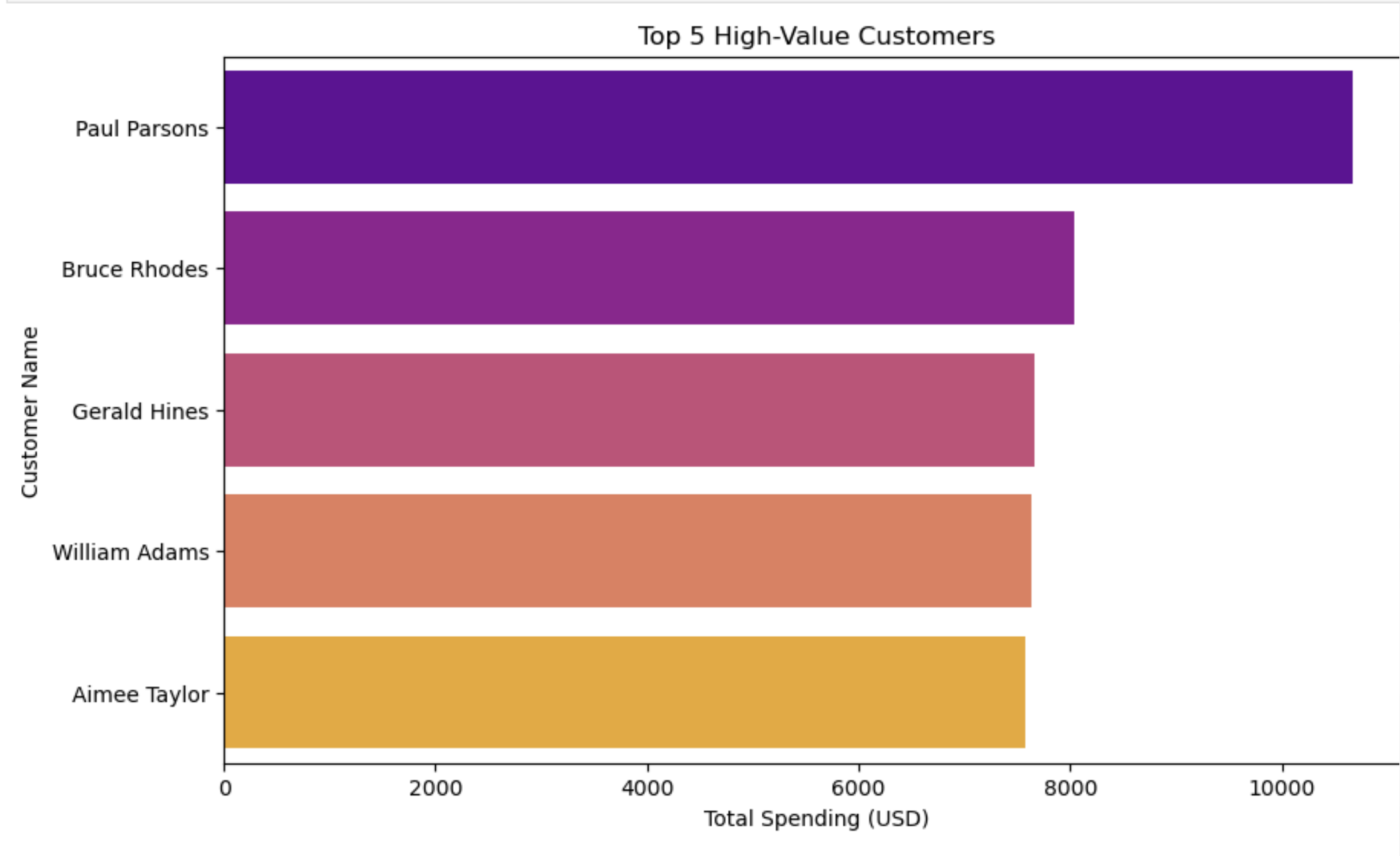
4.PRODUCT CATEGORY DISTRIBUTION

1.THIS PIE CHART SHOWS US WHAT PEOPLE ARE BUYING MOST ON THIS E-COMMERCE SITE.

2."BOOKS" ARE THE BIGGEST HIT, MAKING UP NEARLY 27% OF SALES.

3."ELECTRONICS" AND "HOME DECOR" ARE ALSO POPULAR, WITH AROUND 25% AND 24% OF SALES RESPECTIVELY.

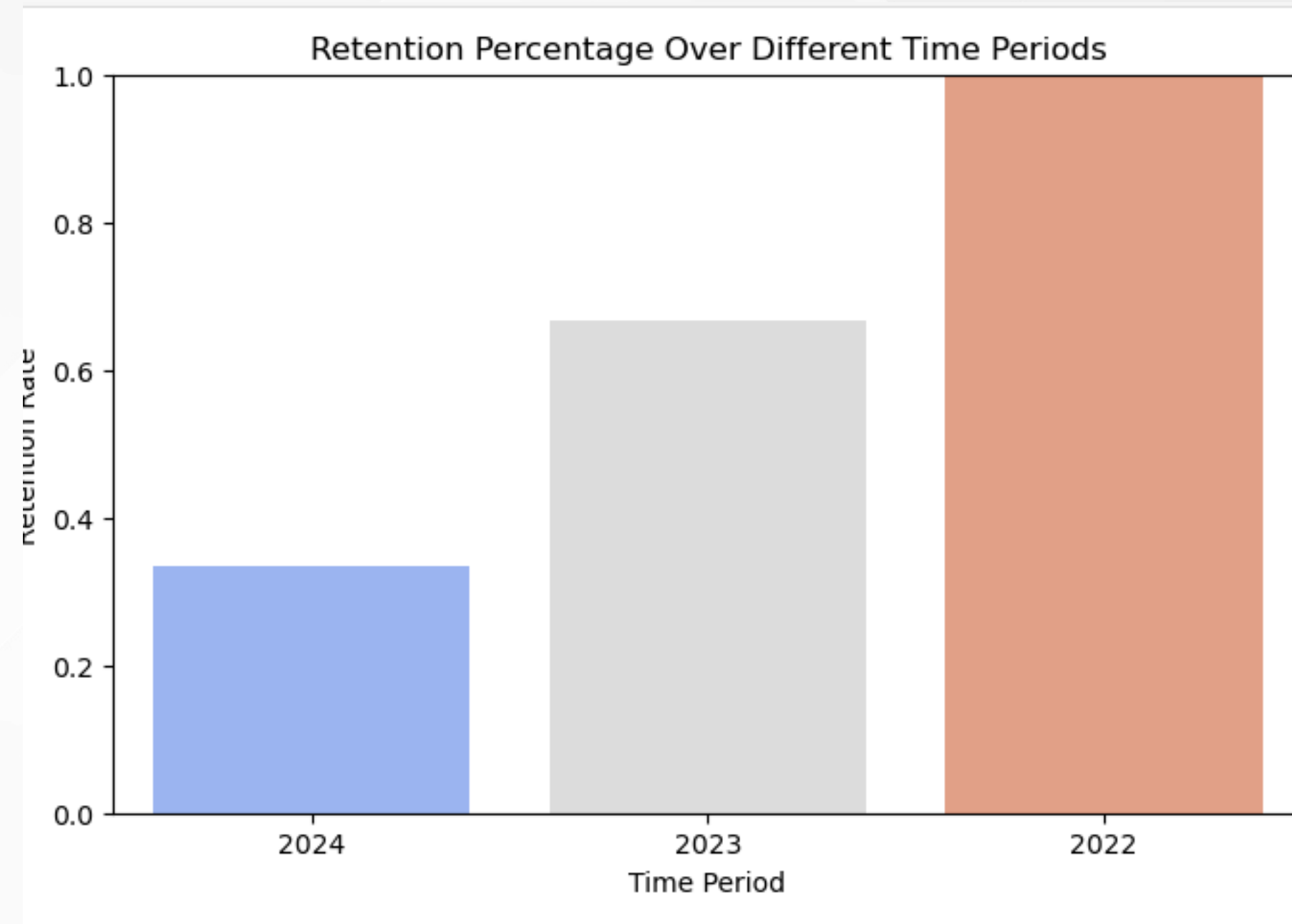4."CLOTHING" IS STILL IMPORTANT, BUT MAKES UP A SLIGHTLY SMALLER PORTION AT 22.8%.

Product Category Distribution

- Books 27.0%
- Electronics 25.4%
- Clothing 22.8%
- Home Decor 24.8%
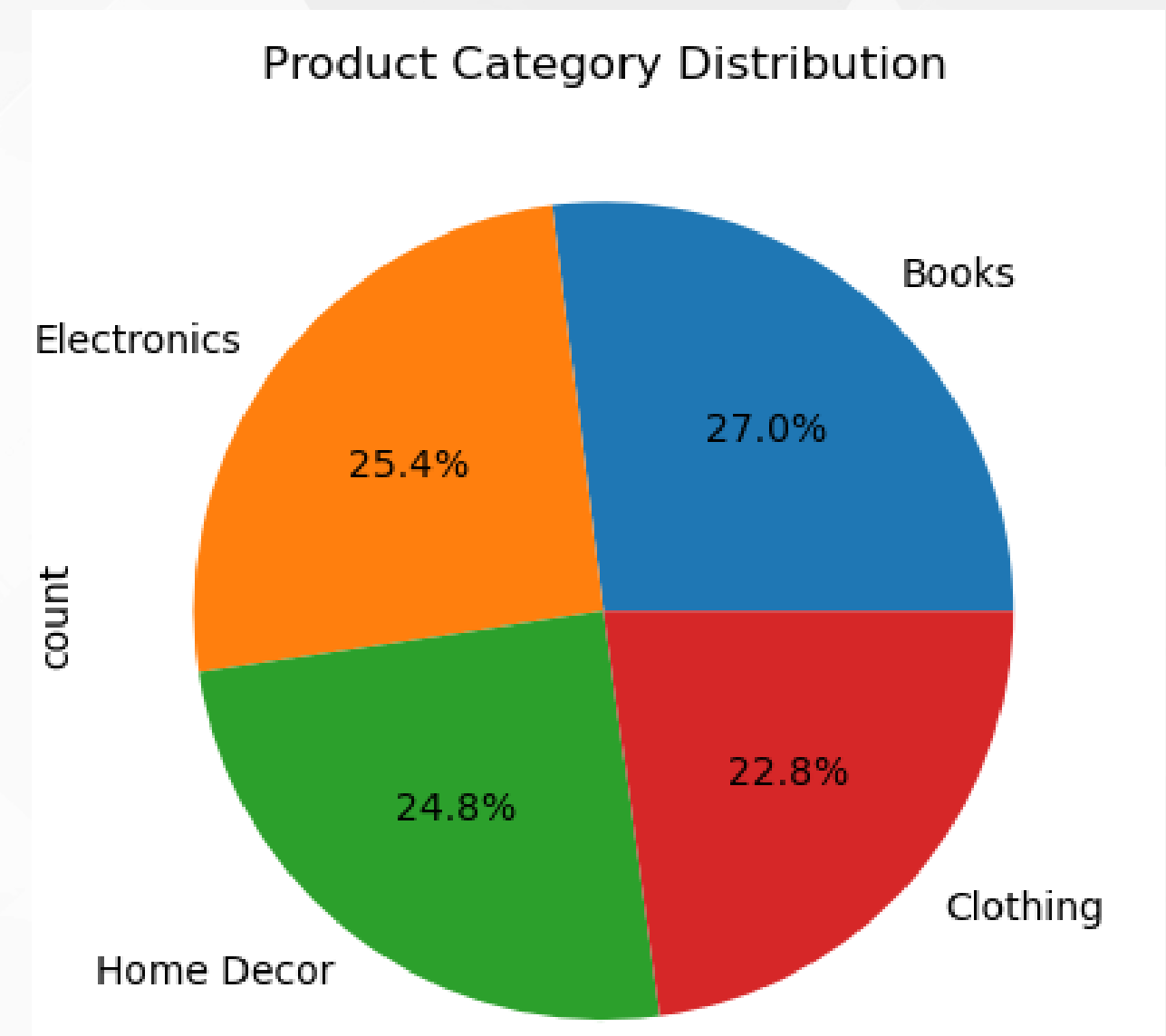
5.SEASONAL TRENDS IN CUSTOMER SIGN-UPS (MONTH-WISE)

- JANUARY SEES THE MOST TRANSACTIONS, LIKELY BECAUSE OF NEW YEAR'S SALES, CLEARANCES, AND PROMOTIONS.

- - THIS IS THE BEST TIME FOR BUSINESSES TO INCREASE REVENUE BY FOCUSING ON STOCK, STAFF, AND MARKETING.

- PREPARING WELL FOR THIS MONTH ENSURES CUSTOMER NEEDS ARE MET, OPERATIONS RUN SMOOTHLY, AND PROFITS SOAR. ALSO, SPECIAL DEALS AND MARKETING CAN FURTHER TAP INTO THIS BUSY PERIOD.

Seasonal Trends in Transactions (Month-Wise)
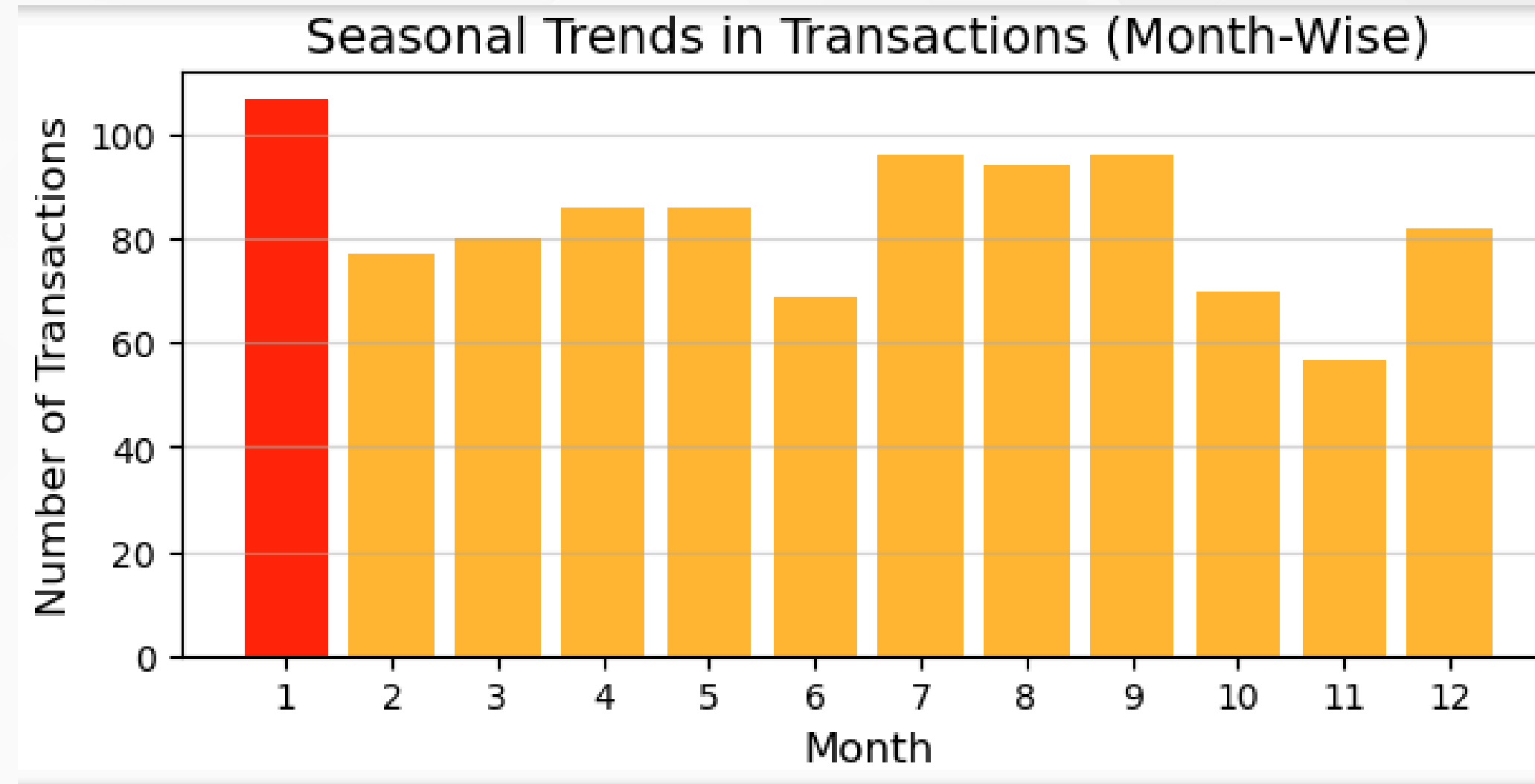
6. RETENTION PERCENTAGE OVER DIFFERENT TIME PERIODS

- RETENTION WAS STRONGEST IN 2022 BUT HAS STEADILY DECLINED THROUGH 2024, SUGGESTING POSSIBLE ISSUES WITH CUSTOMER SATISFACTION OR MARKET CHANGES.

- BUSINESSES SHOULD EXAMINE THE REASONS BEHIND THE HIGH RETENTION IN 2022 AND EXPLORE WHAT LED TO THE DROP AFTERWARD.

- IDENTIFYING THESE FACTORS IS CRUCIAL FOR IMPROVING CUSTOMER LOYALTY AND RETENTION, ENABLING BUSINESSES TO REVERSE THE TREND AND ENHANCE CUSTOMER SATISFACTION IN THE FUTURE.

Retention Percentage Over Different Time Periods

Thank you